

FIG. 1

RNA Isolation Protocol

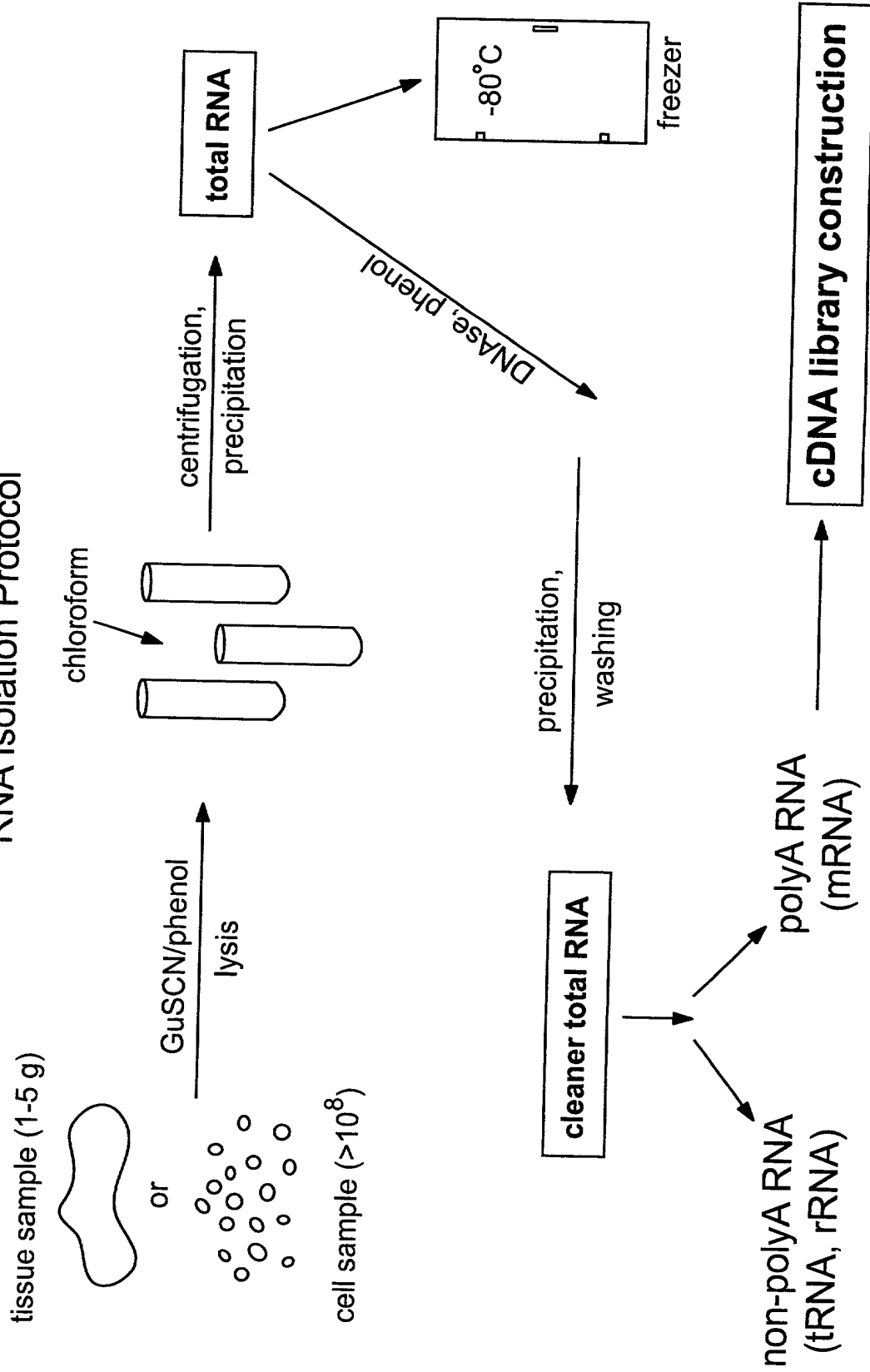
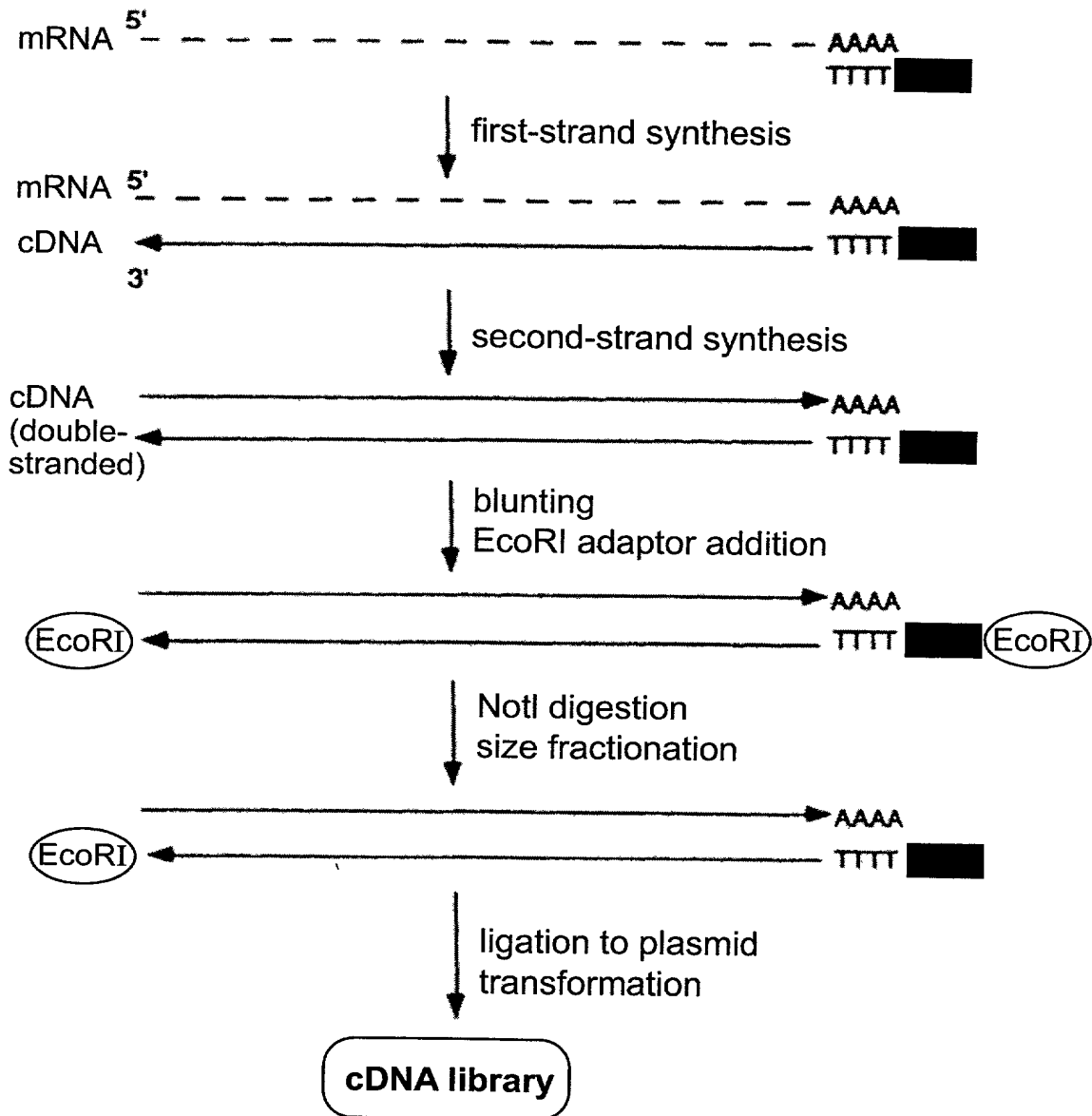


FIG. 2

cDNA Library Construction



Block 1 Sequence Editing Screens

Target Sequence Feature	Editing Method	Result
5' and 3' Vector	Dynamic Programming	Clip
PolyA Tail	Regular Expression	Clip
Sequencing Artifacts	Nearest Neighbor	Remove
Low Information	BLAST (S \geq 90)	Mask
Contamination	BLAST (S \geq 90)	Remove
Repetitive Elements	BLAST (S \geq 90)	Mask
Mitochondrial	BLAST (S \geq 90)	Remove
Ribosomal RNA	BLAST (S \geq 90)	Remove

FIG. 3

FIG. 4

Aberrant				Expected			
A	C	G	T	A	C	G	T
32	8	8	8	14	14	14	14
8	32	8	8	14	14	14	14
8	8	32	8	14	14	14	14
8	8	8	32	14	14	14	14
56	56	56	56	56	56	56	224

e.g. CCCCCGGGTTTTCCCCAAAGGG...

FIG. 5

Rigorous Statistics:

$$E = KNe^{-\lambda S} \text{ where } S = \sum_{HSP} s(a_i, b_j)$$

$$s(a,a) = 5 \text{ and } s(a,b) = -4 \text{ for } a \neq b$$

FIG. 6

Query	Match	BLAST Score	% ID	Length	Product Score
1969935 (238)	1798270	1103	99	259	91
	2057694	765	100	234	65
	1961765	600	100	245	50
	473298	318	92	261	25
473298 (261)	2057694	634	93	234	50
	1961765	634	93	254	46
	1969935	318	92	238	25
	1798270	254	94	259	18

Alignments

1969935 _____
 1798270 _____
 2057694 _____
 1961765 _____
 473298 _____

FIG. 7

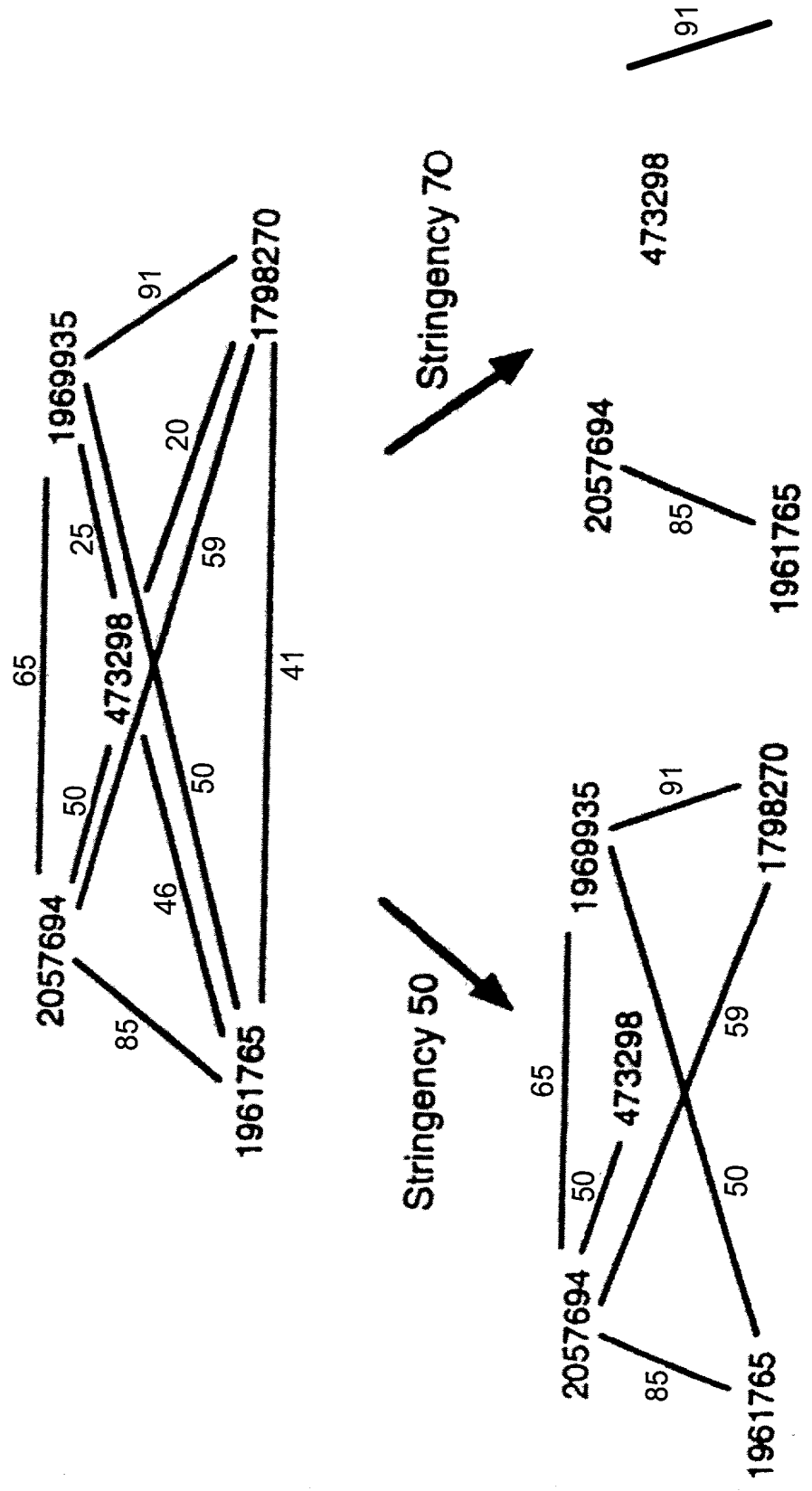


FIG. 8

Creating a Master Cluster

Example: Cluster - 12 (singleton), Cluster 2, and Cluster 1 all contain representative clones with PS ≥ 40 to Gene X.

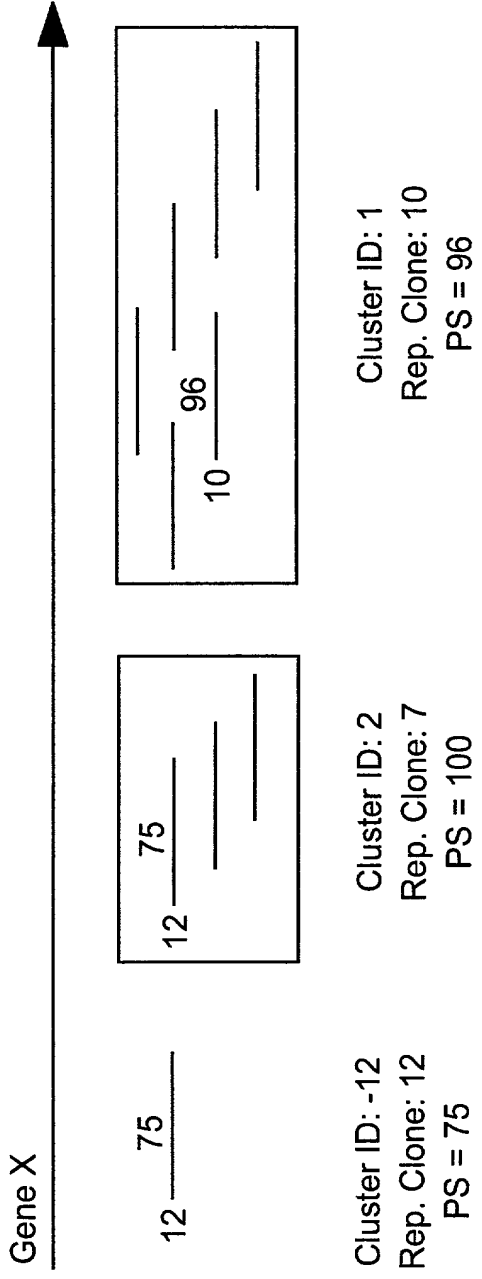
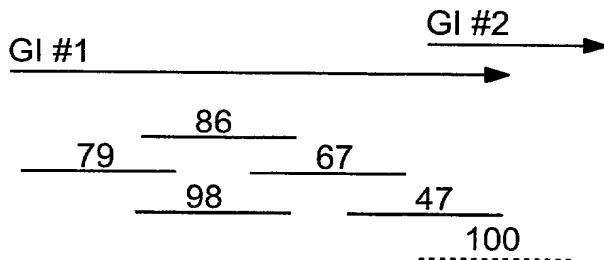


FIG. 9

Naming a Cluster

Q: Does any member of the cluster have a match to GenBank?

YES



5 clones (solid line) match GI #1
 1 clone (dashed) matches GI #2

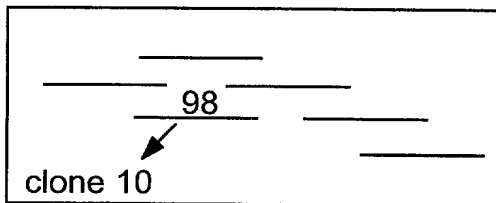


Q: Which GI is represented the most in the cluster?

A: GI #1 is represented the most.

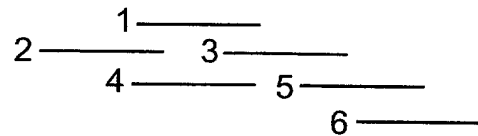


Q: Of those clones matching GI #1, which has the highest Product Score?

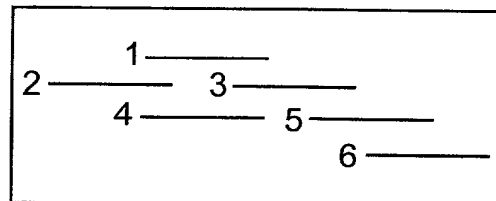


A: Clone 10 has a Product Score of 98, so the cluster is named after this representative clone.

NO



Q: Which clone has the lowest Clone ID for this cluster?



unique cluster

A: Clone 1 has the lowest ID, so the unique cluster is named after this representative clone.

FIG. 10

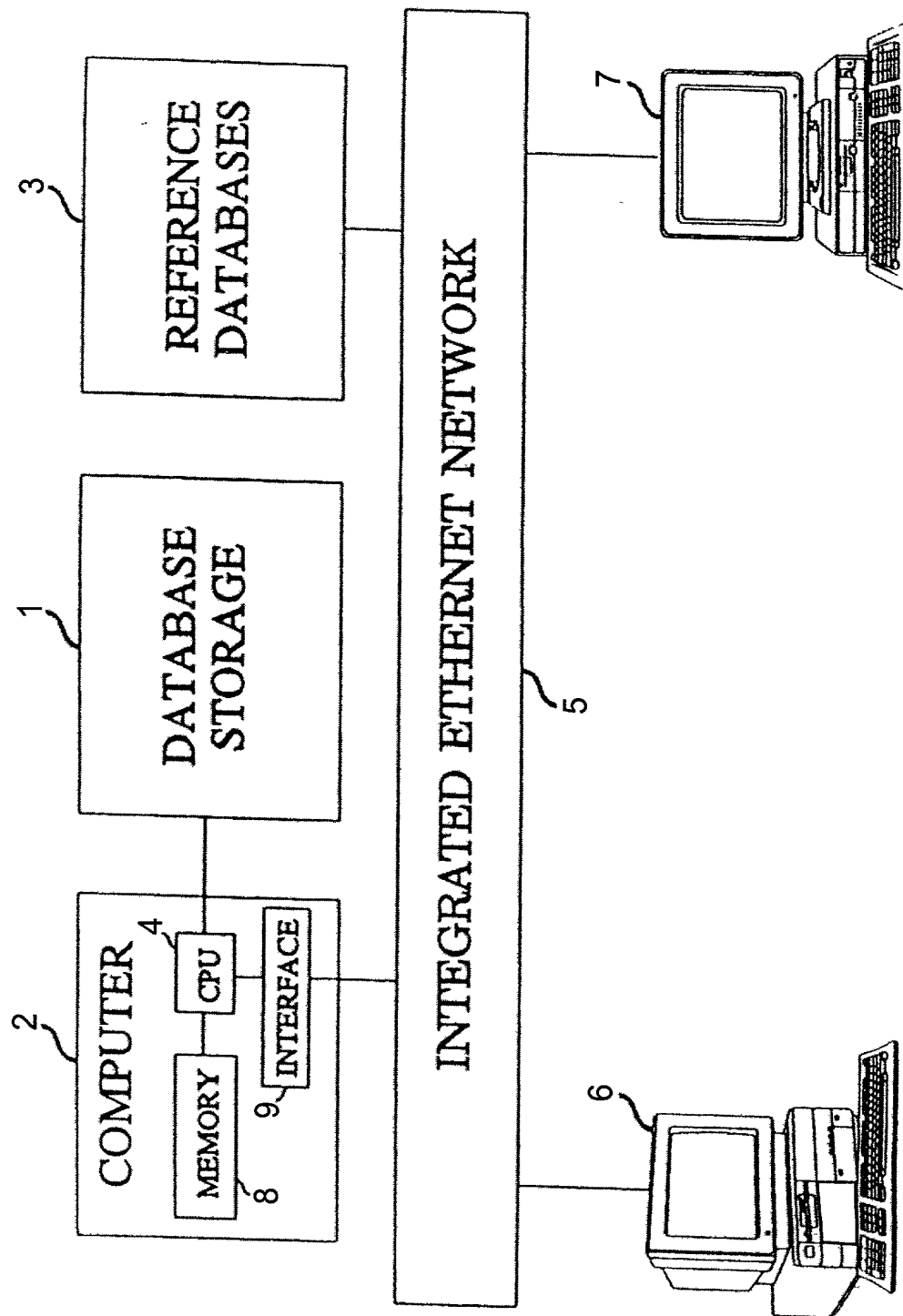


FIG. 11a

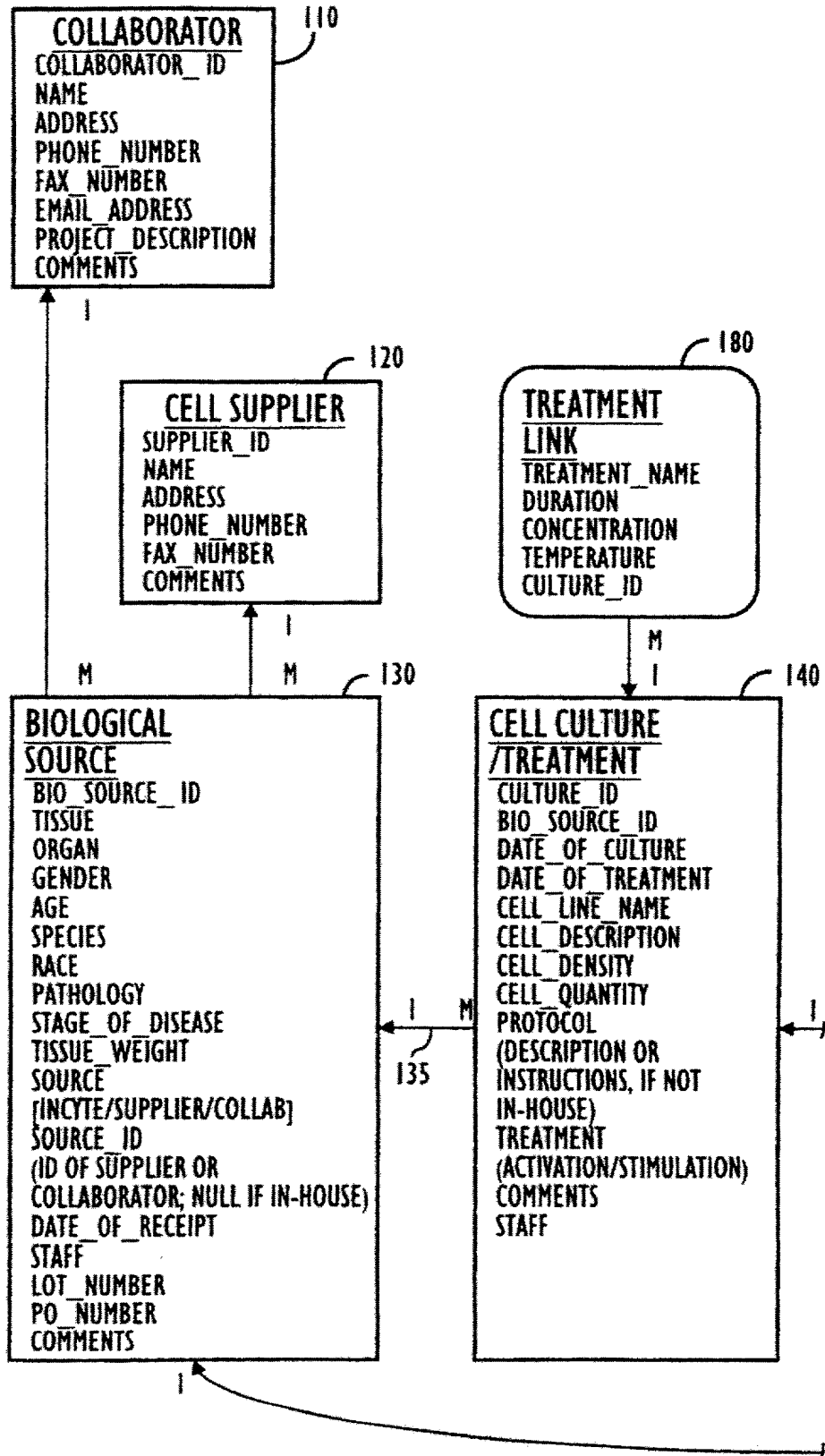


FIG. 11b

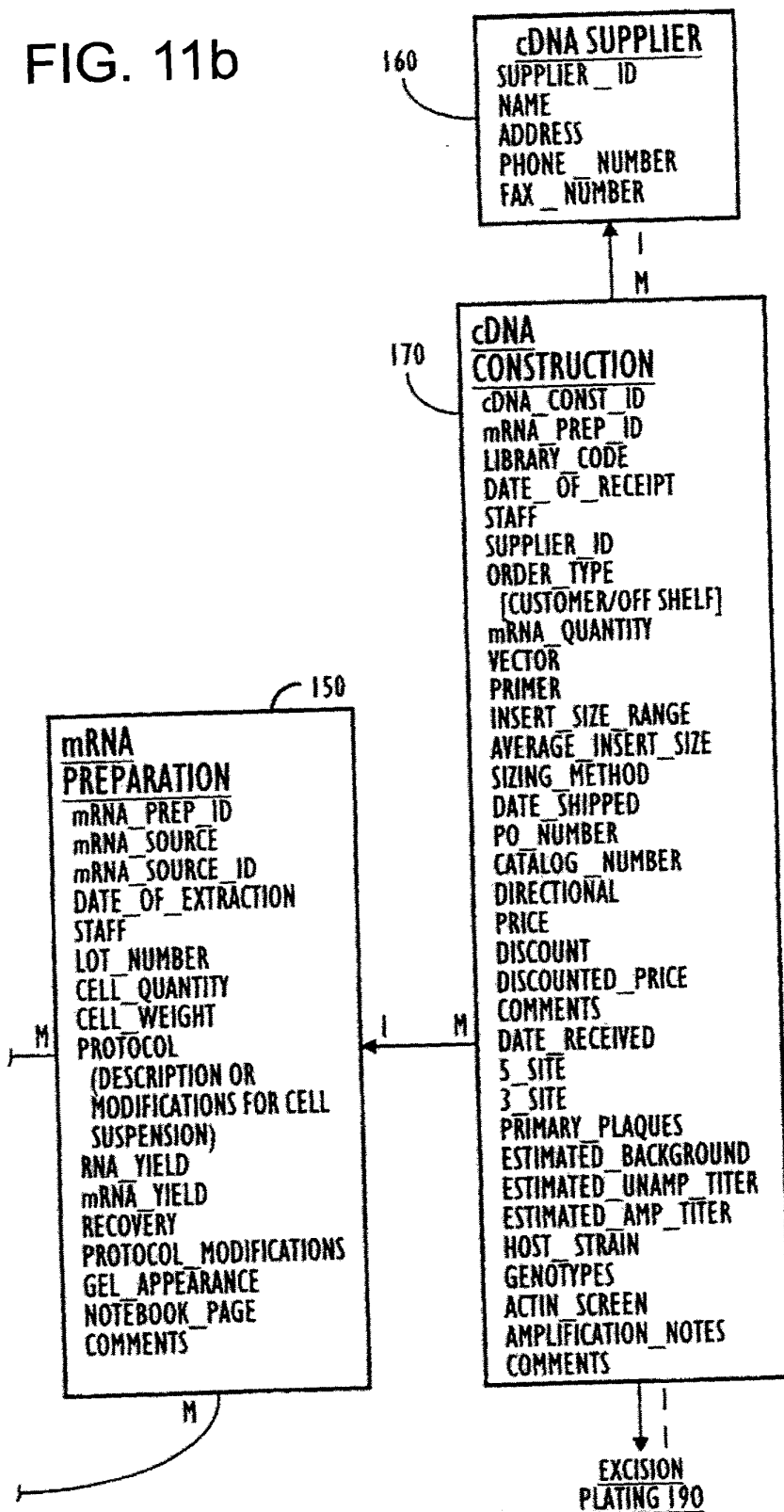


FIG. 12

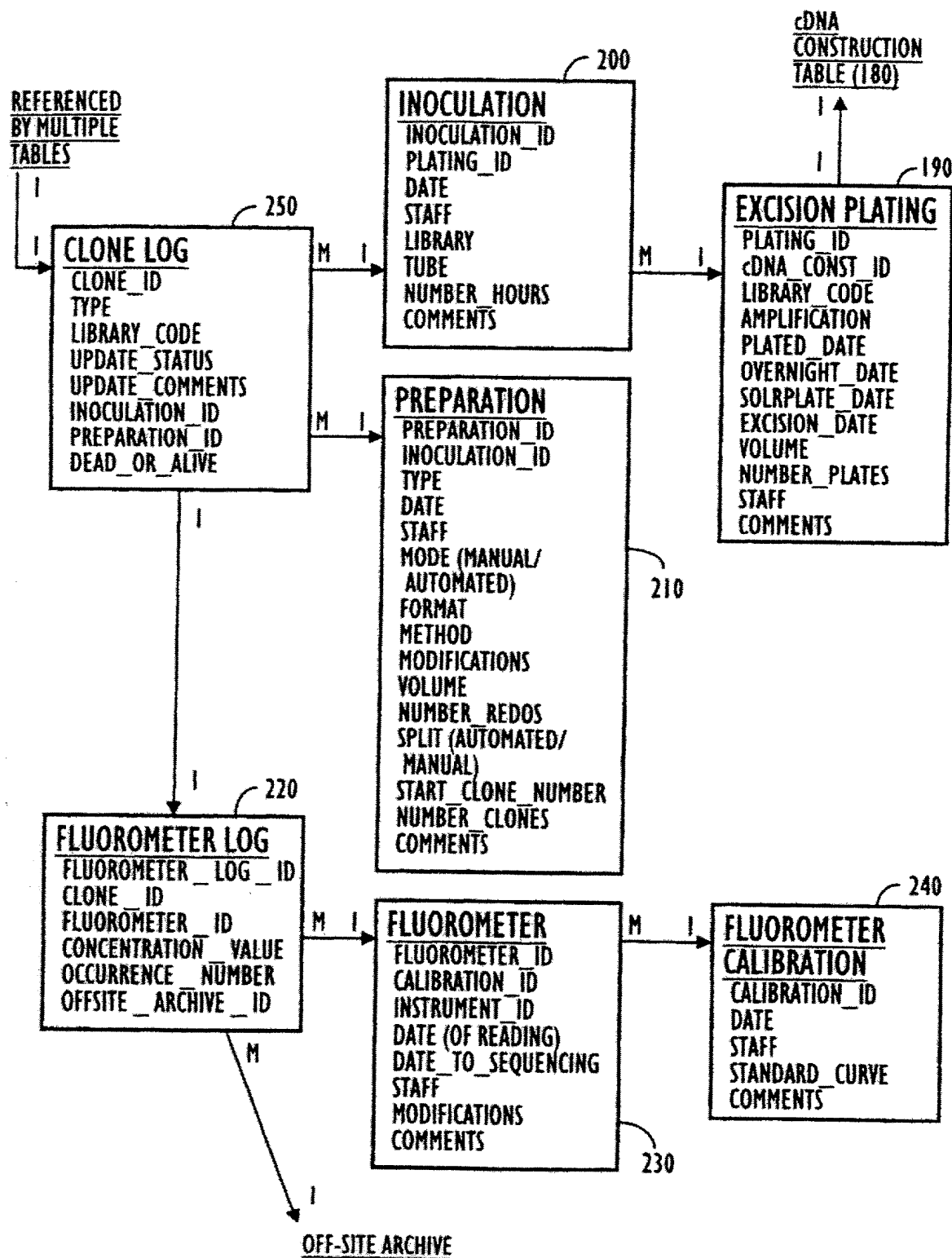


FIG. 13

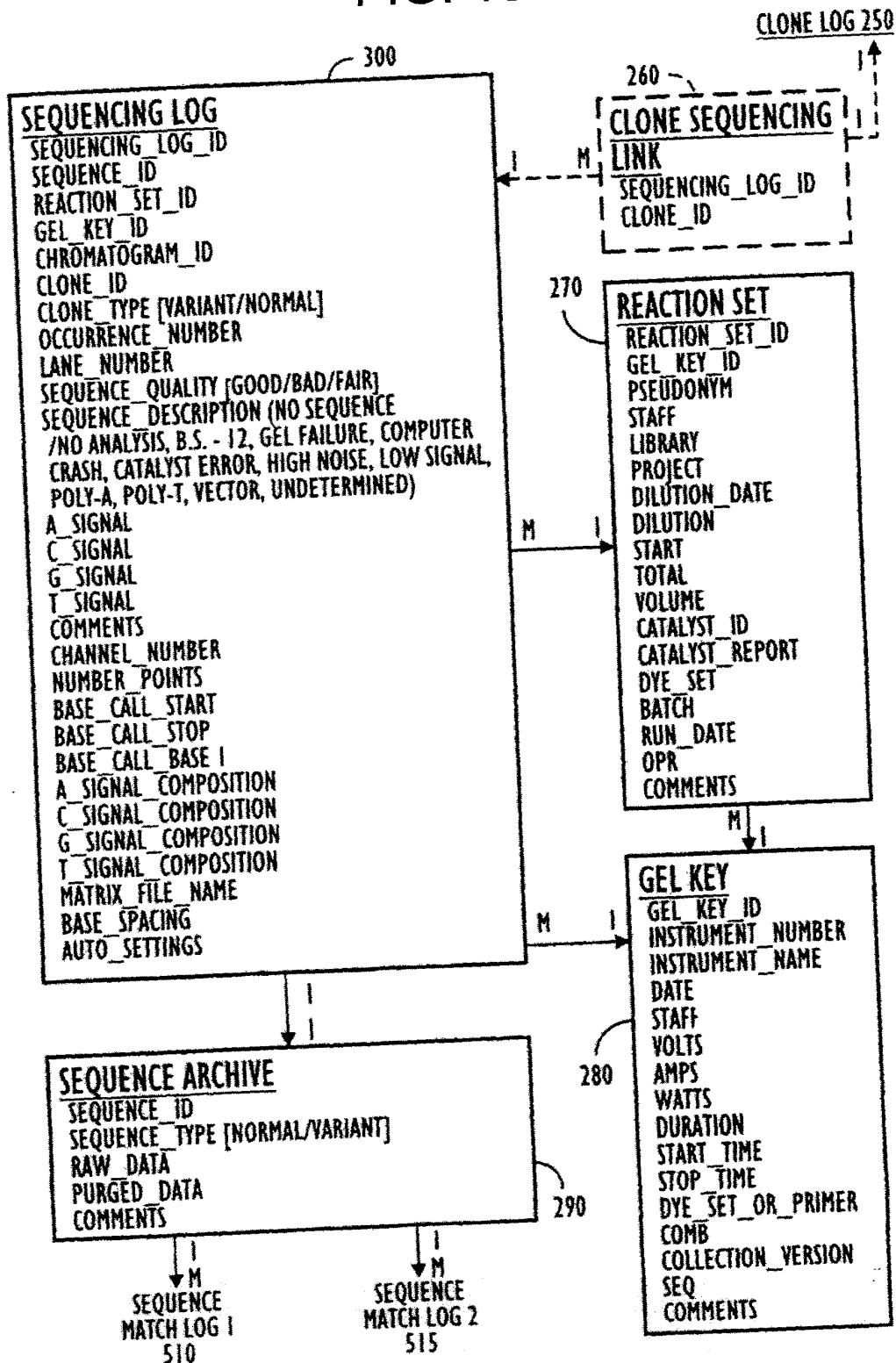


FIG. 14

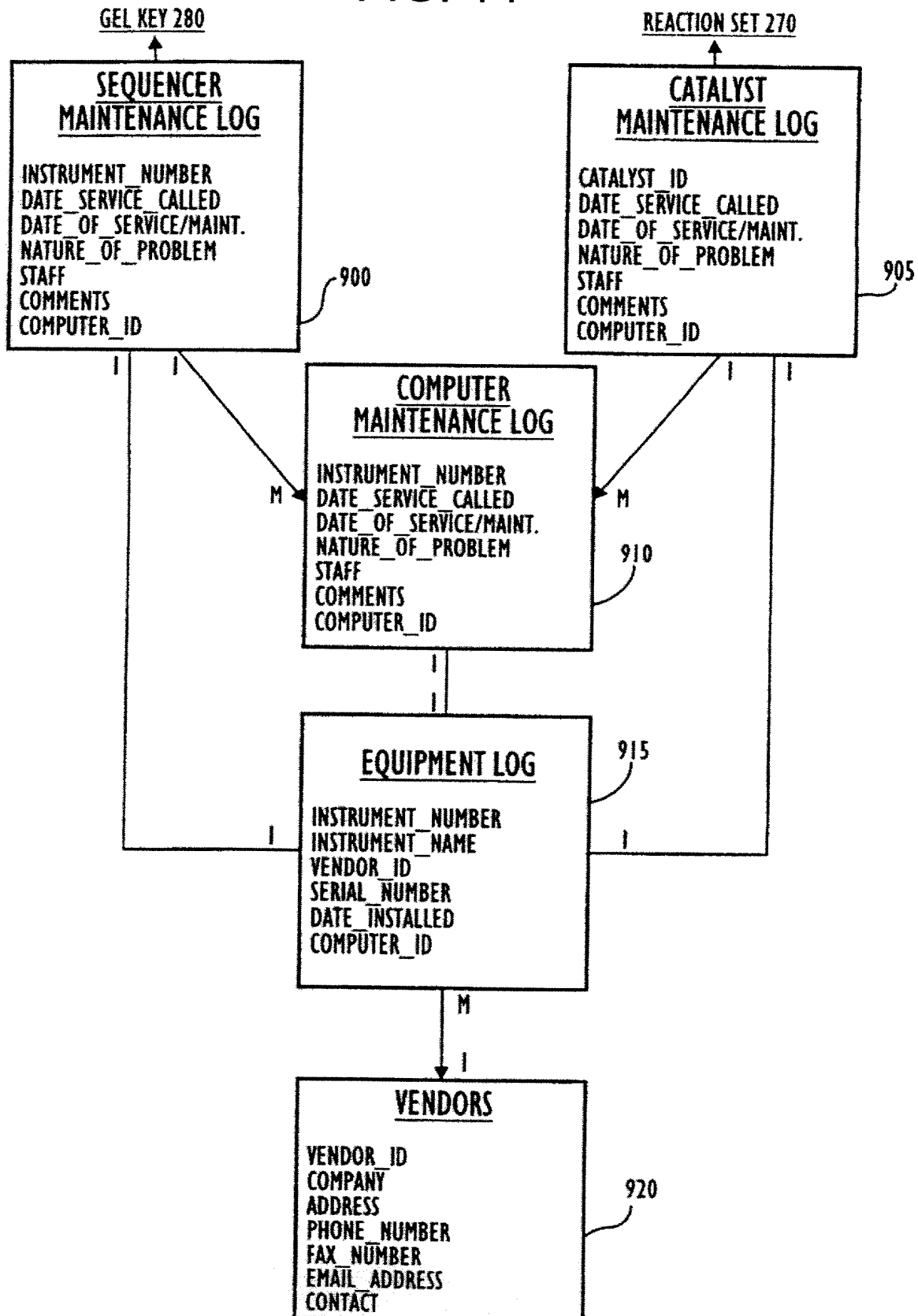


FIG. 15

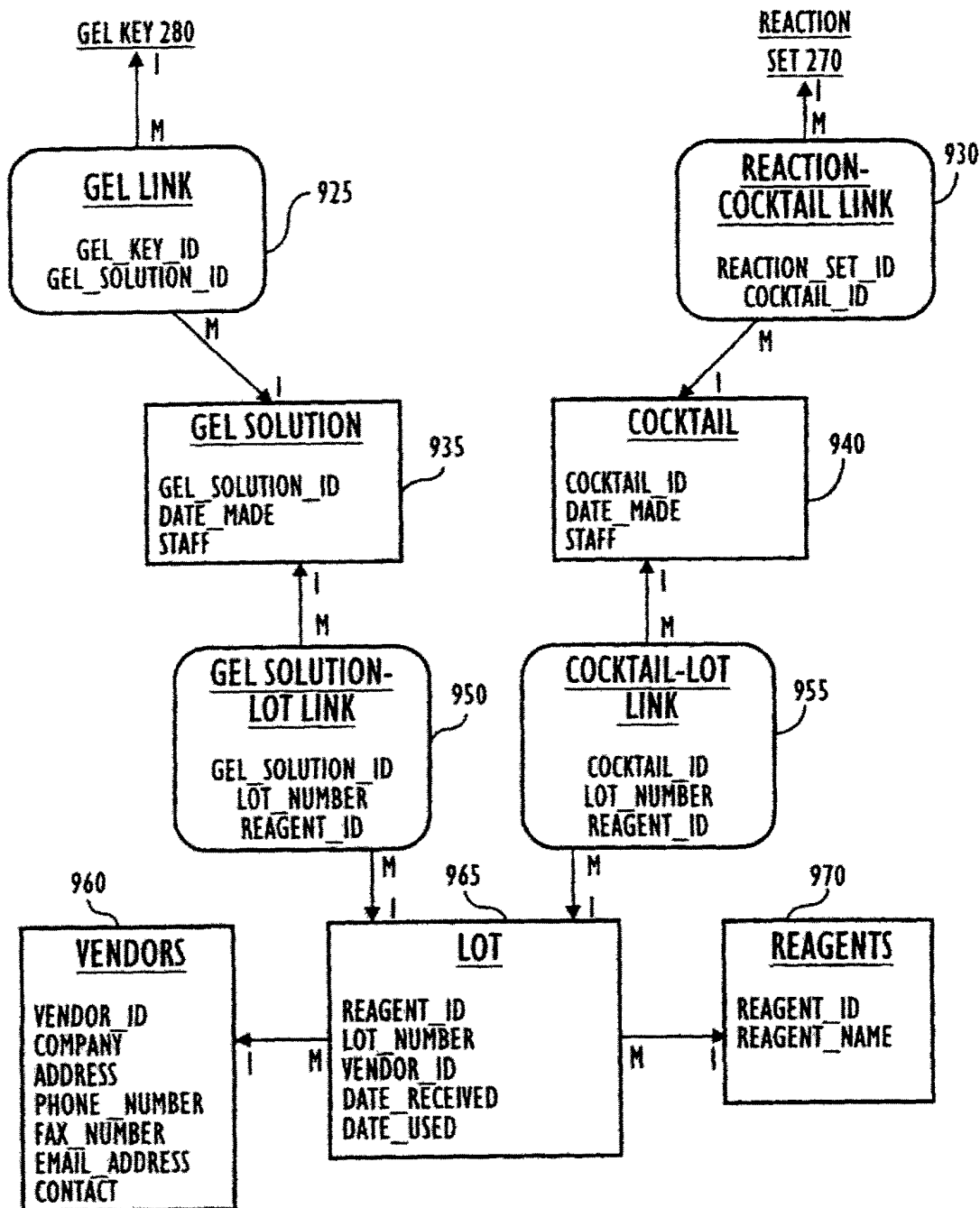


FIG. 16

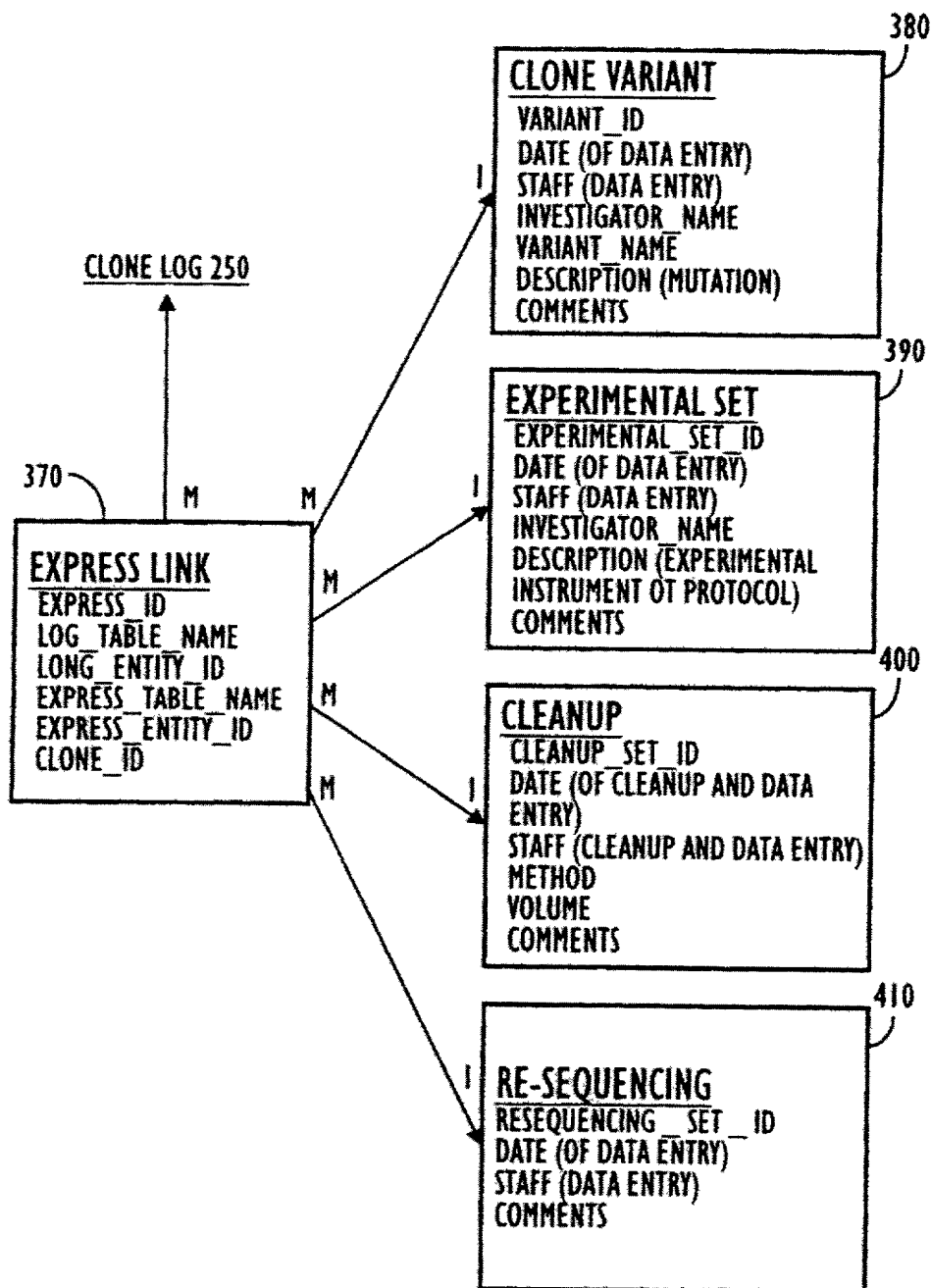


FIG. 18

